XML DATA ENCODING SPECIFICATION FOR INTELLIGENCE PUBLICATIONS VERSION 2 (PUBS.XML.V2)

ICTechSpec 500.D.3-V2

An Intelligence Community Technical Specification
Prepared by the
Intelligence Community Chief Information Officer

24 December 2009

Table of Contents

List of Figures

List of Tables

Chapter 1 – Introduction

1.1 Purpose

This *XML Data Encoding Specification for Intelligence Publications* (PUBS.XML) defines detailed specifications for using Extensible Markup Language (XML) to encode publication metadata in compliance with the Abstract Data Definition. This Data Encoding Specification (DES) defines the XML elements and attributes, associated structures and relationships, mandatory and cardinality requirements, and permissible values for representing the publications concepts using XML. It is the umbrella XML standard for text-based intelligence products.

1.2 Needs and Requirements

This DES is designed to fulfill a number of requirements in support of the transformational efforts of the IC. These requirements include:

- Improving publication and dissemination efficiency by reducing the cost and time for performing manual and complex rendering, manipulation, and content transformation of information resource metadata in context of an intelligence publication.
- Facilitating discovery and exchange of content consisting of mostly text supplemented by interspersed non-textual content (i.e., multi-media information between collectors, all-source analysts, and consumers.
- Capturing an intelligence publication's overall security marking metadata in order to support attribute- and clearance-based information management practices, such as: secure collaboration; content management; content and portion-level filtering of discovery results, and content transfers across security domains.
- Capturing source reference citations to provide intelligence collectors the ability to systematically analyze how and how often the data they gather or produce is being used in order to facilitate better management of collection and production resources.
- Capturing source reference citations to enhance the analytic integrity of formally disseminated intelligence information and improving the traceability of collected information to analytic judgments and conclusions.
- Capturing and retaining a greater understanding of an intelligence publication's meaning, purpose, genesis, and characteristics as identified by a human or service.

1.3 Version Information

This is version 2 of this DES. This version number must be specified in any instance document intended to be valid against this version using the attribute DESVersion. A separate DESVersion attribute should be specified for each DES an instance document is claiming compliance with; each of these attributes should be in the namespace specified by each DES.

The following table lists the major changes made to this DES.

Table 1. Data Encoding Specification Change Summary

Change	Artifacts changed	Compatibility Notes
Various changes to documentation	Documentation Schema	Changes were to correct errors. Any system that relied on the incorrect information may need to be modified.
Removed version number from file names.	Schema	Data generation and ingestion systems need to be updated to use the new file names.
Added ability for instance documents to specify the DES version used for the document.	Schema Constraint Rules	Data generation systems need to be updated to use the new feature. Ingestion systems need to be updated to properly handle the new data.
Changed default namespace.	Schema	Data generation systems need to be updated to use the new namespace. Ingestion systems need to be updated to properly handle the new data.
Added support for Revision/Recall identification.	Schema Constraint Rules	Data generation systems should be updated to use the new structures if they need the feature. Ingestion systems need to use the new specification, including schema and constraints rules. Note: Data could have been created that was valid under previous releases that may not be valid under this release.
Updated reference to DDMS to use version 2.0	Schema	Data generation systems should be updated to use the new structures if they need the feature. Ingestion systems need to use the new specification, including schema. Note: Data could have been created that was valid under previous releases that may not be valid under this release.

Change	Artifacts changed	Compatibility Notes
Updated DES to support ICD 206 source citations.	Schema Constraint Rules Controlled Value Enumerations	Data generation systems need to be updated to use the new feature. Ingestion systems need to be updated to properly handle the new data. Note: Data could have been created that was valid under previous releases that may not be valid under this release.
Relaxed XML Schema for PersonalProfileGroup and replaced with constraint rules.	Schema Constraint Rules	Any system relying solely on the XML schema for validation of PersonProfileGroup may need to be modified to ensure that data are processed appropriately.
Specified and/or update values for numerous elements and attribute to allow more specific validation of data.	Schema Constraint Rules Controlled Value Enumerations	All systems can now use the specified values to ensure data are correct. Note: Data could have been created that was valid under previous releases that may not be valid under this release.

1.4 Audience and Applicability

Data Encoding Specifications are intended primarily to be used by those developing tools and services to create, modify, store, exchange, search, display, or further process the type of data being described. The applicability and conditions for when the DES should be used will be found in the intelligence community standards registry when established, and in approved IC policy guidance written specifically to address information sharing objectives in the context of particular intelligence functions, intelligence information, and desired data formats.

1.5 Utility

A Data Encoding Specification specifies how to implement the abstract data elements in the Abstract Data Definition (ADD) in a particular physical form (e.g., data or file format). For example:

- DESs for textual markup formats, such as Extensible Markup Language (XML) and HTML, define markup elements and attributes, their relationships, cardinalities, processing requirements, and use.
- DESs for display formats, such as text and Adobe Portable Document Format (PDF), define text and typographic conventions, cardinalities, processing requirements, and use.
- DESs for application-specific formats, such as Microsoft Word, define document properties, styles, fields, cardinalities, processing requirements, and use.

1.6 Components of this Data Encoding Specification

This document, in addition to being the core, defining document of the DES, contains the developer's guide. This Data Encoding Specification (DES) consists of a number of technical components: Developer's Guide, Extensible Markup Language (XML) schemas, Controlled Vocabulary Enumeration (CVE) in XML and HTML formats, and an interactive XML Schema Guide.

1.6.1 Developer's Guide

The Developer's Guide provides a mapping of the physical XML structures to abstract data elements defined in the Abstract Data Definition. It also references an interactive XML Schema Guide where more detailed information can be found. The Developer's Guide also includes additional constraint rules, which provide additional restrictions (i.e., constraints) on how the data should be structured and encoded.

1.6.2 Additional Developer's Components

- XML schemas are provided in separate DES artifacts
- CVEs are provided in separate DES artifacts

1.6.3 Example files

Example files are provided to illustrate various encoding situations. The revision/recall and source reference citation examples illustrate the situations presented in the former "Best Practices".

1.7 Normative and Informative Components

The XML schemas, CVE values, and the constraint rules are normative for this Data Encoding Specification. This Developer's Guide, the content referenced within the XML Schema Guide, HTML versions of the CVEs, and Example files are informative.

1.8 Technical Encoding Dependencies

This Data Encoding Specification relies on the *XML Data Encoding Specification for Information Security Marking Metadata* (ISM.XML). Value enumerations used for several XML structures are defined in the various CVEs included in this DES.

1.9 Typographic conventions

Certain typography is used throughout the body of this document to convey certain meanings, in particular:

- *Italics* – A title of a referenced work or a specialized or emphasized term.
- Underscore – An abstract data element.
- **Bold** – An XML element or attribute.

Chapter 2 – Development Guidance

This implementation expands and refines the abstract data elements defined in the ADD, resulting in a physical model for intelligence publications. The implementation incorporates physical formats defined in ISM.XML.

2.1 Mapping of Abstract Data Elements to Physical XML Elements

The mapping of abstract data elements from the ADD to the corresponding physical XML structures defined by this Data Encoding Specification is shown in the following tables, which reflect the groupings in the ADD. These mappings are provided for reference only. The complete set of Data Encoding Specification artifacts, both normative and informative, should be consulted.

This mapping and the mappings in other Data Encoding Specifications provide a starting point for the development of automated transformations between formats defined by the Data Encoding Specifications. However, it should be noted that when these transformations are used between formats with different levels of detail, there might be some data loss.

2.1.1 Publication Metadata

1. Publication Structures (Table 2)
2. Section Structures (Table 3)
3. Narrative Structures (Table 4)

Table 2. Publication Structure Abstract Data Elements to Physical XML Elements

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Article	Used for news or journalist reporting and for other publications with little or no front or rear matter.	**/Article** **/Basic**
Briefing	Used for presentation materials.	**/Briefing**
Correspondence	Used for intelligence content taking the form of memos or letters, with appended materials, as appropriate.	**/Correspondence**
Digest	Used for compendia or consolidation of other publications for packaging and delivery.	**/Digest**
Report	Used for publications with extensive front and rear matter, and body matter that is subdivided into parts, chapters, and/or sections.	**/Report** **/AnalyticalPacket**

Table 3. Section Structure Abstract Data Elements to Physical XML Elements

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Section	Generic subdivision of an article, report, or correspondence.	**//Section**
Section Title	Generic section's primary title.	**//Section/Title**
Sidebar	A short article that is substantially parallel to the text of the main report but not directly a part of it.	**//Sidebar**
Amplification	A section with more detailed information than found in an Analysis section, usually at a higher security classification.	**//Amplification**
Analysis	An exposition of an analyst's research.	**//Analysis**
Appendix	A collection of supplementary material usually placed after the main body of writing.	**//Appendix**
Attachments	A section that is appended or attached to a main document, usually a correspondence document.	**//Attachment**
Bibliography	A list of the works referenced in the body of a publication or consulted by the author in its production.	**//Bibliography** **//Bibliography/BibliographyDivision** **//Bibliography/BibliographyEntry**
Distribution List	A series of addresses or routing symbols for distribution of a publication.	**//DistributionList** **//DistributionEntry**
Glossary	A list of often difficult or specialized words with their definitions.	**//Glossary** **//Glossary/GlossaryDivision** **//Glossary/GlossaryEntry**
Index	Index is an alphabetized list of names, places, and/or subjects that facilitates reference to the body of the publication.	**//Index** **//Index/IndexDivision** **//Index/IndexEntry**

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Key Findings	Key conclusions reached after examination or investigation.	**//KeyFindings**
Preface	A Preface is an introductory section offering information about the source of the request for a report, who wrote the report, the source of the information, how the study was conducted, etc. This element does not specifically address the scope of the report.	**//Preface**
Scope	The Scope is the extent or range of application, aim or purpose of a report.	**//Scope**
Summary	A comprehensive and usually brief abstract, recapitulation, or compendium of facts, statements, and/or findings.	**//Summary**
Table of Contents	Listing of sections, figures, tables, or other specially titled content listed by title within the publication and pointer to the content.	**//TOC**

Table 4. Narrative Structure Abstract Data Elements to Physical XML Elements

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Equation	A complex structure representing a formula or an expression, such as a mathematical or chemical equation.	**//Equation**
List	Series of items representing distinct but related thoughts written together in a meaningful grouping or sequence.	**//List** **//List/ListItem** **//AttachmentList** **//AttachmentList/ListItem**
Media Resource	Media Resource is a complex structure including a form of presentable media (e.g., graphic, animation, video) and some form of unique identification (e.g., title) or clarification (e.g., legend).	**//MediaResource**
Note	Comment or explanation further clarifying surrounding content.	**//Note** **//NoteInline**
Paragraph	A distinct portion of written matter dealing with a particular idea.	**//Para**
Quote	Passage copied verbatim from a book, speech or other source that is properly referenced.	**//Quote**
Table	Complex structure including a two-dimensional list organized into a grid containing rows and columns with special presentation characteristics and some form of unique identification (e.g., title).	**//Table**

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Assertion	A complex structure used to highlight content and associate special emphasis (via formatting), semantic understanding (via tagging) or ancillary value-added information (via hyperlink).	**//Account** **//BEnumber** **//CommData** **//Commodity** **//Concept** **//CountryName** **//Date** **//DateTime** **//EntityUntyped** **//Equipment** **//Event** **//Facility** **//GeoFeature** **//GeoRef** **//Identifier** **//InfoBearer** **//LocationOfInterest** **//MilitaryUnit** **//Money** **//Narcotic** **//Nomenclature** **//Organization** **//Person** **//QuantityReference** **//SystemClass** **//Term** **//Time** **//Vehicle** **//Weapon**
Footnote	A note that comments on—or cites a reference for—a designated part of the content, usually presented inline, at the bottom of the page, or at the end of a publication.	**//Footnote**
Source Citation	Bibliographic citation specialized to identify information sources necessary to substantiate analysis.	**//SourceGroup** **//SourceGroup/SourceReference**

2.1.2 Information Resource Metadata

All Xpaths in this table, which fall under the PublicationMetadata element, are prefixed with the corresponding parent element of either **AdministrativeMetadata** or **DescriptiveMetadata**. Xpaths without either of these prefixes, such as **/SourceGroup**, may appear in other contexts in a publication XML file.

Table 5. Information Resource Metadata Conceptual Elements to Physical XML Elements

ICS Concept	ICS Definition	XPath and XML implementation notes
Contributor	An entity responsible for making contributions to the resource. Examples of Contributor include a person, an organization, or a service. Typically, the name of a Contributor should be used to indicate the entity.	**//AdministrativeMetadata/Contributor**
Coverage	The spatial, temporal [or virtual] topic of the resource, the spatial [or virtual] applicability of the resource, or the jurisdiction under which the resource is relevant. Spatial topic may be a named place or a location specified by its geographic coordinates. Temporal period may be a named period, date, or date range. Virtual topic may be a named place or a location specified using a network or email address. A jurisdiction may be a named administrative entity or a geographic place to which the resource applies. Recommended best practice is to use a controlled vocabulary such as the Thesaurus of Geographic Names (TGN) or the NGA GEOnet Names Server (GNS) as sanctioned by the United States Board on Geographic Names. Where appropriate, named places or time periods can be used in preference to numeric identifiers such as sets of coordinates or date ranges.	**//DescriptiveMetadata/Coverage**
Creator	An entity primarily responsible for making the resource. Examples of Creator include a person, an organization, or a service. Typically, the name of a creator should be used to indicate the entity.	**//AdministrativeMetadata/Publisher**

ICS Concept	ICS Definition	XPath and XML implementation notes
Date	A point or period of time associated with an event in the lifecycle of the resource. Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3CDTF profile of ISO 8601. Typically, date will be associated with the creation or availability of the resource.	**//AdministrativeMetadata/DateList**
Description	An account of the resource. Description may include but is not limited to: an abstract, a table of contents, a graphical representation, or a free-text account of the resource.	**//DescriptiveMetadata/Description**
Format	The file format, physical medium, or dimensions of the resource. Examples of dimensions include size and duration. Recommended best practice is to use a controlled vocabulary such as the list of Internet Media Types (MIME). Format may be used to identify the software, hardware, or other equipment needed to display or operate the resource.	**//AdministrativeMetadata/RecordsManagementInfo**
Identifier	An unambiguous reference to the resource within a given context. Recommended best practice is to identify the resource by means of a string conforming to a formal identification system. Formal identification systems include but are not limited to the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI), and the International Standard Book Number (ISBN).	**//AdministrativeMetadata/IdentifierList**
Language	A language of the resource. Recommended best practice is to use a controlled vocabulary such as RFC 3066, *Tags for the Identification of Languages*, which specifies use of ISO 639-2, *Codes for the Representation of Names of Languages*, three character language code, with an optional appended ISO 3166-1, *Codes for the representation of names of countries and their subdivisions*, two character country code. For example: "eng-US" or "eng-UK."	**//DescriptiveMetadata/Language**

ICS Concept	ICS Definition	XPath and XML implementation notes
Publisher	An entity responsible for making the resource available. Examples of a Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity.	**//AdministrativeMetadata/Publisher**
Relation	A related resource. Recommended best practice is to identify the referenced resource by means of a label or number conforming to a formal identification system.	**//AdministrativeMetadata/Relation**
Rights	Information about rights held in and over the resource. Typically, rights will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the rights element is absent, no assumptions may be made about any rights held in or over the resource.	**//AdministrativeMetadata/Rights** **//MediaResource/CopyrightAttribution**
Resource Security Mark	The overall security classification and security handling instructions carried by the resource. Resource Security Mark applies to the resource-level classification, SCI controls, dissemination controls, non-IC markings, and other security provisions prescribed by Executive Order 12958, as amended, the Information Security Oversight Office (ISOO) Directive 1 of the National Archives and Records Administration, and the Intelligence Community marking standard maintained by the Controlled Access Program Coordination Office (CAPCO). These values are prominently presented, in the case of intelligence publications, at the top and bottom of every page and in other specified locations. See the *Intelligence Community Standard for Information Security Marking Metadata* for refinements of this conceptual element.	**//DescriptiveMetadata/Security** **//Attachment/Security**
Source	The resource from which the described resource is derived. The described resource may be derived from the related resource in whole or in part. Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system.	**//SourceGroup**

ICS Concept	ICS Definition	XPath and XML implementation notes
Subject	A topic of the resource. Typically, the topic will be represented using keywords, key phrases, or classification codes. Recommended best practice is to use a controlled vocabulary. To describe the spatial, temporal or virtual topic of the resource, use the Coverage element.	**//DescriptiveMetadata/Activity** **//DescriptiveMetadata/Subject**
Title	A name given to the resource. Typically, a Title will be a name by which the resource is formally known.	**//DescriptiveMetadata/AbbreviatedTitle** **//DescriptiveMetadata/Title** **//DescriptiveMetadata/Subtitle**
Type	The nature or genre of the content of the resource. The Type includes terms describing general categories, functions, genres, or aggregation levels for content. Examples of Types include publication forms (e.g., reports or articles) and intelligence disciplines (e.g., SIGINT, MASINT, HUMINT). Recommended best practice is to use a controlled vocabulary. To describe the file format, physical medium, or dimensions of the resource, use the Format element.	**//DescriptiveMetadata/IntelType** **//DescriptiveMetadata/Note** **//OtherProperty** **//DescriptiveMetadata/PositiveIntel** **//DescriptiveMetadata/ProductLine** **/Article, /Briefing, /Correspondence, /Digest, /Report**

2.1.3 Source Reference Citation Metadata

The conceptual elements defined in the ICS are mapped to XML elements as shown in Table 6. Conceptual element refinements are mapped to XML elements as shown in Table 7. In the mapping, classification information is usually mapped to the Information Security Markup ISM and denoted as @ism:* to represent the collection of attributes defined by ISM.

Table 6. Source Reference Citation Abstract Data Elements to Physical XML Elements

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Bibliographic Resource (DCMI)	A book, article, or other documentary resource. **IC Expansion:** In the context of source reference citations, a bibliographic resource is all significant and substantive reporting or other information upon which analytic judgment, assessments, estimates, or confidence levels depend. An intelligence product may be derived from one or more source references in whole or in part. Recommended best practice is to identify a related information resource by means of a formal identification system.	An actual document being cited.
Bibliographic Citation (DCMI)	A bibliographic reference for the [cited] resource. **IC Expansion:** A special type of bibliographic reference (i.e., a formal identification system) unique to the intelligence discipline that contains pertinent information resource metadata and details of the extent of the information being referenced. In accordance with ICD 206, source reference citations are to be listed in a special section at the end of intelligence products.	**//SourceReference** Except as otherwise specified, all XPath specifications in Table 7 are relative to **//SourceReference**.

Table 7. Source Reference Citation Abstract Data Element Refinements to Physical XML Elements

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Citation Security Mark	Classification marking used for the overall Bibliographic Citation. This is the citation's portion mark as displayed in the bibliography or collection of source references.	**@ism:***
Consulted	A date and time when a cited resource was used as a basis for analytic judgment.	**DateTimeReferenced**
Creator (DCMI)	An entity primarily responsible for making the [cited] resource. **IC Expansion:** The Creator can represent an author and/or coauthor and/or point of contact for the cited resource. The entity must be from or associated with the originating organization defined by the Publisher. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the Citation Security Mark should be followed.	**POCinfo** – IC originator **AuthorInfo** – Human author(s) **EditorInfo** – Editor(s) Classification information is provided in: **@ism:*** within the appropriate element.
Date of Information	A date, time range, or time period representing the relative currency of the specific information cited.	**DateInformation**
Identifier (DCMI)	An unambiguous reference to the [cited] resource within a given context. **IC Expansion:** Recommended best practice is to identify a related information resource by means of a formal identification system. Examples might include a report serial number, document name or number, image frame identification code, or an organization internal identification or tracking number.	**DocumentID** – Unique resource ID **SourceID** – Agency internal identifier or tracking number
Issued (DCMI)	Date of formal issuance (e.g., publication) of the [cited] resource.	**DatePublished**
Link	A hyperlink to the cited resource. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the Citation Security Mark should be followed.	**Link/@xlink:href** or **SegmentReferenced/Link/@xlink:href** Classification information is provided in: **Link/@ism:***

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Publisher (DCMI)	An entity responsible for making the [cited] resource available. **IC Expansion:** An IC element, national government, international organization, or open-source owner(s) and/or producer(s) of a cited resource. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the Citation Security Mark should be followed.	IC originator: **AgencyAcronym** Open source: **POCinfo** Foreign government: **CountryCode** and optionally **POCinfo** The type of publisher (IC, open-source, or foreign), is provided in: **@publisherType** Classification information is provided in: **AgencyAcronym/@ism:*** or **CountryCode/@ism:*** or **POCinfo/@ism:***
Segment Referenced	An identifier or description of the extent of the cited resource. Typically includes a form of label (e.g., a section or paragraph number, image feature, page number or range, video frame or range, etc.), possibly the classification of the extent, and possibly a link into the cited resource. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the Citation Security Mark should be followed.	**SegmentReferenced/MediaExtent** or **SegmentReferenced/SegmentLabel** or **SegmentReferenced/Link/@xlink:href** Classification information is provided in: **SegmentReferenced/ReferencedResourceSecurity/@ism:*** The security attributes should represent the contemporary CAPCO marks for the segment referenced. The security attributes on the ReferencedResourceSecurity element do NOT impact the overall security mark of the citation since they only convey what the classification of the segment referenced is not any data in the current document. Additionally if the segment referenced has a legacy CAPCO or other marking the text form of that marking could be put in **SegmentReferenced/ReferencedResourceSecurity/OriginalClassificationMarking** Should the **OriginalClassification** text be in and of itself classified that Classification information may be provided in: **SegmentReferenced/ReferencedResourceSecurity/OriginalClassificationMarking/@ism:***

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Sourced Content	A word, phrase, sentence, or other contiguous text string for which attribution is being cited. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the Citation Security Mark should be followed.	**//SourceGroup/SourcedText** Classification information is provided in: **//SourceGroup/SourcedText/@ism:***
Source Descriptor	An explanation of factors contained in the cited resource or publicly available information that the producing organization assesses may affect the quality or reliability of the information in the specific cited resource. Factors may include, but are not necessarily limited to, completeness, precision or technical quality, context, or age/currency of the information. In the case of human sources, this explanation may include information that describes the level of access, past reporting record, or potential biases (e.g., political, personal, professional, or religious affiliations). If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the Citation Security Mark should be followed.	**SourceDescriptor** Classification information is provided in: **SourceDescriptor/@ism:***
Source Security Mark	Overall classification marking of the cited resource. As the resource could originate from the US or another country, the Source Security Mark should represent an appropriate US marking or an original non-US marking. The originating country of the classification marking should also be recorded.	**ReferencedResourceSecurity** The security attributes should represent the contemporary CAPCO marks for the cited resource. The security attributes on the Security element do NOT impact the overall security mark of the citation since they only convey what the classification of the cited resource is not any data in the current document. Additionally if the cited resource has a legacy CAPCO or other marking the text form of that marking could be put in **OriginalClassificationMarking** Should the **OriginalClassification** text be in and of itself classified that Classification information may be provided in: **OriginalClassificationMarking/@ism:***

Abstract Data Element	Abstract Data Definition	XPath and XML implementation notes
Title (DCMI)	A primary title of the [cited] resource. **IC Expansion:** There may be multiple titles associated with a given cited resource, especially if the resource is published as part of a larger compilation of materials. Titles associated with a publication, journal, series, or edition may be necessary to uniquely indentify the cited resource. Recommended best practice is to provide the cited resource's title, possibly an alternative title if one exists, and to provide additional titles of the larger compilations when necessary. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the Citation Security Mark should be followed.	**Title** **Subtitle** **CompilationTitle** **EditionNumber** **VolumeNumber** **IssueNumber** Classification information is provided in: **Title/@ism:*** **Subtitle/@ism:*** **CompilationTitle/@ism:***
Type (DCMI)	The nature or genre of the content of the [cited] resource. **IC Expansion:** The Type includes terms describing general categories, functions, genres, or aggregation levels for content. Examples of Type include publication form (e.g., book, periodical, report, or article), online publication (e.g., Internet site, web page, blog, or wiki), or intelligence discipline (e.g., SIGINT, MASINT, HUMINT). Recommended best practice is to use a controlled vocabulary. If applicable, data associated with this concept should be classification marked and appropriate rules for displaying the marking or for influencing the value or display of the Citation Security Mark should be followed.	**SourceType** **IntelDiscipline** Should the **IntelDiscipline** be classified it would impact the classification of the overall source citation by changing: **IntelDiscipline/@ism:***

2.1.4 Information Security Marking Metadata

The PUBS.XML schemas incorporate the information security marking attributes defined by the *XML Data Encoding Specification for Information Security Marking Metadata* (ISM.XML). These XML attributes are associated with XML elements throughout the model; most often wherever a portion or information resource-level security banner or classification/declassification block is displayed, but also on many other XML elements that may never have their markings displayed.

2.2 Additional Guidance

This section contains descriptions of how particular sitations, which may be encountered in data to be encoding using this Data Encoding Specification, should be handled. In particular, situations that should be addressed using multiple elements or attributes should be documented here. The content of this section will evolve over time as additional sitations are identified.

2.2.1 Specification of Publishing Organization

The element **Publisher** is used to identify the entity(ies) primarily responsible for releasing the information to the enterprise. The entity(ies) of interest in this context are foremost the organization responsible; however, the organizations and/or individuals responsible for creating the information are also captured within the **Publisher** structure. The publishing organization's approved identifier value is captured in a child element called **AgencyAcronym**. The organizations and/or individuals responsible for creating the information are captured in child elements called **AuthorInfo**, **CoauthorInfo**, **ContributingAuthorInfo**, and **POCInfo**. Depending on the enterprise requirement being addressed, a complete understanding of the **Publisher** requires evaluating the **AgencyAcronym/@acronym** value as well as the values found in the **Affiliation** and **OfficeName** structures of the **AuthorInfo**, **CoauthorInfo**, **ContributingAuthorInfo**, and **POCInfo** elements.

The **Publisher** structure provides the ability to identify multiple levels of organizational structure and multiple organizations or individuals responsible for creating the information. The most basic ability to identify a publisher uses the required **AgencyAcronym** element. The default controlled vocabulary enumeration (CVE) for **AgencyAcronym** includes values representing the organizations officially designated as part of the IC as defined in the DNI's *Overview of the United States Intelligence Community for the 111th Congress* of 2009, plus the DNI, plus additional entries intended to recognize non-IC publishers whose information is commonly used in support of the intelligence mission. One of these values must be selected.

In many cases, the AgencyAcronym CVE only includes the highest level of the organization structure (e.g., DNI), service or agency (e.g., US Army, DHS, DoS), or non-IC designation (e.g., OtherDoD, Foreign). In order to identify a Publisher at a level below what the AgencyAcronym CVE allows, use the OfficeName element of the AuthorInfo or POCInfo structures. The **OfficeName** element is officially defined as the name or organizational code of a department, division, branch, or other office within an agency.

For consistency, populate OfficeName with an approved organization acronym designator for the sub-organization. For multiple levels of sub-organization, list the

acronyms in descending order delimited with the "/" character. As one or both of **AuthorInfo** and **POCInfo** are required, the **POCInfo/OfficeName** will have precedent over the **AuthorInfo/OfficeName**, if both are included.

In cases where non-IC information (e.g., OtherDoD, OtherUSG, SLT, Foreign) is shared with the intelligence enterprise, the **AgencyAcronym** should reflect the organization, which last prepared the information for consumption (e.g., converted the content into PUBS.XML, applied enhanced information resource metadata tagging, translated, or packaged the information into an official IC product) and shared the product with the enterprise. As that organization is affecting the record status of the product, it must take responsibility for addressing any questions about the information.

If a non-IC producer is providing information that is already compliant with IC enterprise data encoding standards, then the AgencyAcronym should reflect the appropriate non-IC organization designator and the non-IC organizations office in the appropriate **OfficeName** element. Examples of this scenario might exist in a USG department where there are sub-organizations designated in the IC and sub-organizations not in the IC; DoD where some sub-organizations support DIA, some support a service, and some are not in the IC; State, Local, Tribal organizations with information that flows into the intelligence enterprise via DHS, NCTC, or other means; or with our foreign partners. In the case of foreign partners designations in the **OfficeName**, precede the office acronym with the country code tri-graph in order to ensure uniqueness.

Examples

For NCTC:

```
<Publisher><AgencyAcronym acronym=”DNI”>
<POCInfo><OfficeName>NCTC</OfficeName></POCInfo></Publisher>
```

For the XYZ component of NCTC:

```
<Publisher><AgencyAcronym acronym=”DNI”>
<POCInfo><OfficeName>NCTC/XYZ</OfficeName></POCInfo></Publisher>
```

For the XYZ component of CIA:

```
<Publisher><AgencyAcronym acronym=”CIA”>
<POCInfo><OfficeName>XYZ</OfficeName></POCInfo></Publisher>
```

For the United State Postal Service:

```
<Publisher><AgencyAcronym acronym=”OtherUSG”>
<POCInfo><OfficeName>USPS</OfficeName></POCInfo></Publisher>
```

For the JIOC at PACOM:

```
<Publisher><AgencyAcronym acronym=”DIA”>
<POCInfo><OfficeName>PACOM/JIOC</OfficeName></POCInfo></Publisher>
```

For the J4 at PACOM:

```
<Publisher><AgencyAcronym acronym=”OtherDoD”>
<POCInfo><OfficeName>PACOM/J4</OfficeName></POCInfo></Publisher>
```

Chapter 3 – XML Schema Guide

The detailed description and reference documentation for the PUBS.XML schema can be found in a separate document entitled *PUBS.XML Schema Guide*. This guide serves as an interactive presentation of the PUBS.XML schema as well as an implementation-specific data element dictionary.

The guide was generated with a commercially available product named *XML Spy®,* produced by Altova. The physical XML structures illustrated in the guide are described in Appendix H.

The guide provides an interactive index to:

- Global Elements and Attributes,
- Local Elements and Attributes,
- Simple and Complex Types,
- Groups and Attribute Groups, and
- Referenced Schemas.

Where applicable, the guide provides:

- Diagram,
- Namespace,
- Type,
- Children,
- Used by,
- Properties,
- Patterns,
- Enumerations,
- Attributes,
- Annotations, and
- Source Code.

The guide is published in a folder consisting of a master HTML file (SgPubsXml.html), with supporting graphics and HTML files.

Chapter 4 – Data Validation Constraint Rules

The constraint rules in this chapter define data validation constraints for PUBS.XML beyond those in the XML Schema. The rules are written in plain English phrases; however, knowledge of the PUBS.XML is required to understand the rules. The constraint rules will eventually be offered in a declarative form, such as Schematron.

4.1 Basics

The PUBS.XML schema defines the data elements, attributes, cardinalities and parent-child relationships for which XML instances must comply. Validation of these syntax aspects is an important first step in the validation process. An additional level of validation is needed to ensure that the content complies with the constraints as specified by IC elements and codified in these constraint rules. Traditional schema languages are generally unable to effectively represent these additional constraints. The constraints can largely be derived from interpreting the intent of the current schemas and/or extracting guidance from the Developer's Guides or the policy or guidance documents that govern that type of data.

4.1.1 Terminology

For the purposes of this document, the following statements apply:

- The term "is specified" indicates that an attribute is applied to an element and the attribute has a non-null value.
- The term "must be specified" indicates that an attribute must be applied to an element and the attribute must have a non-null value.
- The term "is not specified" indicates that an attribute is not applied to an element, or an attribute is applied to an element and the attribute has a null value.
- The term "must not be specified" indicates that an attribute must not be applied to an element.

4.1.2 Levels of Validation

The constraint rules defined in this document are described by levels of constraint. While the remainder of this document provides *specific* guidance for the constraints at each level for PUBS.XML, the general intent of each constraint level is defined below:

- *Well-Formed.* A document is Well-Formed if it meets the W3C's definition of "well-formed".

- *Schema Valid.* A document is considered Schema Valid if it is Well-Formed and conforms to all of the rules embodied in the referenced schema.
- *Constraint Rules Valid.* A document is considered Constraint Rules Valid if it is Schema Valid and conforms to the specific constraint rules defined for the referenced schema.

The World Wide Web Consortium (W3C) defines "Well-Formed"[1] in the following way:

A textual object is a *Well-Formed* XML document if

- Taken as a whole, it matches the production labeled document.
- It meets all the well-formedness constraints given in the specification.
- Each of the parsed entities, which are referenced directly or indirectly within the document, is well-formed.

In accordance with the W3C XML recommendation, an XML document is considered Schema Valid if it is Well-Formed and conforms to all of the parent-child, cardinality, datatype, and namespace rules embodied in the referenced schema.

As is evident by these definitions, schema and constraint rule validations can be performed only if a document references a schema.

4.1.3 Rule Identifiers

Each constraint rule has an assigned rule ID, indicated in brackets preceding the constraint rule description. The rule IDs from 00001 to 10000 are unclassified and 10001 to 20000 are "for official use only" (FOUO). IDs from 20001 to 30000 are reserved for Secret rules and 30001 and above for more classified rules. PUBS.XML data validation constrain rule IDs are prefixed with "PUBS-ID-"

As the constraint rules are managed over time, IDs from deleted rules will not be reused.

4.1.4 Errors and Warnings

The severity of a constraint rule violation is categorized as either an "Error" or a "Warning." An "Error" is more severe and is indicative of a clear violation of a constraint rule, which would be likely to have a significant impact on the quality of a document. A "Warning" is less severe although noteworthy, and may not necessarily have any impact

[1] Extensible Markup Language (XML) 1.0 (Third Edition), http://www.w3.org/TR/2004/REC-xml-20040204/

on the quality of a document. The severity of a constraint rule violation is indicated in brackets preceding each constraint rule description

Each system responsible for processing a document (e.g., create, modify, transform, or exchange) must make a mission-appropriate decision about using a document with errors or warnings based on mission needs.

4.2 Non-null Constraints

XML syntax allows all elements with content declared to be of datatype "string" to have zero or more characters of content — which, allows for empty (or null) content. According to this Specification, all required elements (and certain conditional elements) must have content, other than white space. If an element, defined in this Specification, used in an XML instance is required (or conditional in certain cases), and that element may possibly contain only text content, then the element must have content in order to be Constraint Rules Valid.

[PUBS-ID-00001][Warning] For any optional element that exists and can have text content, the element should have a non-null, non-whitespace value.

[PUBS-ID-00002][Warning] For every optional attribute that exists in the document, a non-whitespace value must be specified.

[PUBS-ID-00003][Error] The following elements, if they exist in an XML instance, must have content:

- **TaskID**
- **DescriptiveMetadata/Title**
- **Language**
- **DescriptiveMetadata/Description**
- **Keyword**
- **MediaFormat**
- **ApplicationSoftware**
- **NetworkAddress**
- **Appendix/Title**
- **Attachment/Title**
- **BibliographyDivision/Title**
- **BibliographyEntry**
- **Definition/Para**
- **DistributionEntry**
- **GlossaryDivision/Title**
- **GlossaryTerm**
- **IndexDivision/Title**

- **IndexEntry**
- **ListItem**
- **LongDescription/Para**
- **Section/Title**
- **Briefing/Slide/Title**
- **ArticleGroup/Title**
- **NIPFTopic**

[PUBS-ID-00004][Error] If element **NIPFAssertion** exists, at least one instance of child element **CountryCode** or child element **NonStateActorCode** must have content.

[PUBS-ID-00005][Error] Replaced by [PUBS-ID-00065]

[PUBS-ID-00006][Error] Replaced by [PUBS-ID-00065]

[PUBS-ID-00007][Error] If element **GeographicIdentifier** exists, at least one of its child elements must have content.

- **GeographicIdentifier/CountryCode**
- **GeographicIdentifier/Facility**
- **GeographicIdentifier/PlaceName**
- **GeographicIdentifier/Region**

[PUBS-ID-00008][Error] The following elements must have element content in at least one of their child elements, excluding the **PublicationMetadata** child element.

- **AnalyticalPacket**
- **Article**
- **Basic**
- **Briefing**
- **Correspondence**
- **Report**

[PUBS-ID-00009][Error] For element **SourceReference**, either child element **SourceID** or child element **DocumentID** must exist and must have a non-null value.

4.3 DES Constraints

The DES version is specified through attributes on the root element. These attributes are constrained by the following rules. The DESVersion enables systems processing an instance document to be certain which set of schema, CVE's and business rules are intended by the author to be used.

[PUBS-ID-00077][Error] The root element must have the attribute **DESVersion** in the namespace urn:us:gov:ic:ism with the value of "2".

[PUBS-ID-00078][Error] The root element must have the attribute **DESVersion** in the namespace urn:us:gov:ic:pubs with the value of "2".

4.4 Value Enumeration Constraints

Several elements and attributes of the PUBS.XML model use Controlled Vocabulary enumerations (CVEs) to define the data allowed in the element or attribute. In some cases the specific CVE is specified via an attribute, which may include a default CVEs. Further, in some of the cases where the CVE can be specified, the attribute may restrict the list of CVEs allowed and some may allow for the author to specify their own CVE. For each of these, the value must be in the specified external CVE or the default CVE.

Some CVEs are not available on all networks. A subset CVE will be provided for use on networks not approved for the entire list. An example of such a case is CVEnumPubsAgencyAcronym where the non-FOUO values are provided in CVEnumPubsAgencyAcronym.xml on non-FOUO networks. If the processing will occur on a network where the entire CVE is not available, the subset CVE may be substituted in the constraint rules since the excluded values would be excluded from use on the lower network.

As noted in the specific rules below, a failure of validation against a CVE will generate an Error.

[PUBS-ID-00010] Replaced by PUBS-ID-00066, PUBS-ID-00067, PUBS-ID-00068, and PUBS-ID-00069.

[PUBS-ID-00066][Error] If element **CountryCode** has attribute **vocabulary** specified as "FIPS" the element value must be in CVEnumPubsCoverageFIPSDigraph.xml.

[PUBS-ID-00067][Error] If element **CountryCode** has attribute **vocabulary** specified as "ISO-2" the element value must be in CVEnumPubsCoverageISO3166Digraph.xml.

[PUBS-ID-00068][Error] If element **CountryCode** has attribute **vocabulary** specified as "ISO-3" the element value must be in CVEnumPubsCoverageISO3166Trigraph.xml.

[PUBS-ID-00069][Error] If element **CountryCode** has attribute **vocabulary** specified as "ISO-NR" the element value must be in CVEnumPubsCoverageISO3166Numeric.xml.

[PUBS-ID-00011][Error] If element **CountryCode** does not have attribute **vocabulary** specified the element value must be in CVEnumPubsCoverageFIPSDigraph.xml.

[PUBS-ID-00012]Replaced by PUBS-ID-00071, PUBS-ID-00072, and PUBS-ID-00073.

[PUBS-ID-00071][Error] If element **CountryName** has attribute **countryCodeVocabulary** specified as "FIPS" the attribute **countryCode** value must be in CVEnumPubsCountryFIPSDigraph.xml

[PUBS-ID-00072][Error] If element **CountryName** has attribute **countryCodeVocabulary** specified as "ISO-3" the attribute **countryCode** value must be in CVEnumPubsCountryISO3166Trigraph.xml

[PUBS-ID-00073][Error] If element **CountryName** has attribute **countryCodeVocabulary** specified as "ISO-NR" the attribute **countryCode** value must be in CVEnumPubsCountryISO3166Numeric.xml

[PUBS-ID-00013]Replaced by PUBS-ID-00056, PUBS-ID-00057, PUBS-ID-00058, PUBS-ID-00059, PUBS-ID-00060, and PUBS-ID-00061

[PUBS-ID-00056][Error] If element **Language** has the attribute **encoding** value of "ISO639-1" then the **Language** element value must be in CVEnumISO639Digraph.xml and no country code portion may be specified in the **Language** element value.

[PUBS-ID-00057][Error] If element **Language** has the attribute **encoding** value of "ISO639-2" then the **Language** element value must be in CVEnumISO639-2Trigraph.xml and no country code portion may be specified in the **Language** element value.

[PUBS-ID-00058][Error] If element **Language** has the attribute **encoding** value of "ISO639-3" then the **Language** element value must be in CVEnumISO639-3Trigraph.xml and no country code portion may be specified in the **Language** element value.

[PUBS-ID-00059][Error] If element **Language** has the attribute **encoding** value of "RFC1766" then the language code portion of the **Language** element value must be in CVEnumISO639Digraph.xml and the country code portion, if present, must be in CVEnumPubsCoverageISO3166Digraph.xml.

[PUBS-ID-00060][Error] If element **Language** has the attribute **encoding** value of "RFC3066" then the language code portion of the **Language** element value must be in CVEnumISO639Digraph.xml or CVEnumISO639-2Trigraph.xml and the country code portion, if present, must be in CVEnumPubsCoverageISO3166Digraph.xml.

[PUBS-ID-00061][Error] If element **Language** has the attribute **encoding** value of "RFC4646" then the language code portion of the **Language** element value must be in

CVEnumISO639Digraph.xml or CVEnumISO639-2Trigraph.xml and the country code portion, if present, must be in CVEnumPubsCoverageISO3166Digraph.xml.

[PUBS-ID-00014][Error] Replaced by PUBS-ID-00062, PUBS-ID-00063, and PUBS-ID-00064.

[PUBS-ID-00062][Error] If the element **QuantityReference has the** attribute **unitOfMeasureVocabulary** with a value of "UNECE20" the value of attribute **unitOfMeasure** must be in CVEnumUNCE20UnitsOfMeasure.xml.

[PUBS-ID-00063][Error] If the element **QuantityReference** has the attribute **unitOfMeasureVocabulary** with a value of "ISO4217-3" the value of attribute **unitOfMeasure** must be in CVEnumISO4217Trigraph.xml.

[PUBS-ID-00064][Error] If the element **QuantityReference** has the attribute **unitOfMeasureVocabulary** with a value of "ISO4217-NR" the value of attribute **unitOfMeasure** must be in CVEnumISO4217Numeric.xml.

[PUBS-ID-00074][Error] If the element **MediaFormat** is specified the element value must be in CVEnumPubsMimeType.xml.

[PUBS-ID-00075][Error] If the attribute **MIMEtype** is specified the value must be in CVEnumPubsMimeType.xml.

[PUBS-ID-00076][Error] If the element **AgencyAcronym** does not have the attribute **vocabulary** specified the value of attribute **acronym** must be in CVEnumPubsAgencyAcronym.xml.

4.5 Revision Recall Constraints

[PUBS-ID-00070][Error] Element **OtherProperty** must not have the attribute **vocabulary** with a value of "RevisionRecallDesignations".
Human Readable: The Revision Recall Best practice method of using **OtherProperty** is now incorporated into the schema using **RevisionRecall** therefore use in this DES is rescinded.

[PUBS-ID-00079][Error] Element **DescriptiveMetadata/Title** must not start with any of the following strings "ADMINISTRATIVE RECALL", or "ADMINISTRATIVE REVISION", or "SUBSTANTIVE RECALL", or "SUBSTANTIVE REVISION".
Human Readable: The Revision Recall Best practice method of using text in the **Title** is now incorporated into the schema using **RevisionRecall** therefore use in this DES is rescinded.

4.6 Element Based Constraints

[PUBS-ID-00015][Error] If element **CountryCode** has attribute **precedence** with a value of "Secondary," there must be at least one sibling element **CountryCode** for which attribute precedence has a value of "Primary."

[PUBS-ID-00016][Error] All required attributes in a PUBS.XML instance must have a value.

[PUBS-ID-00017][Error] If element **CountryName** has either attribute **countryCode** or **countryCodeVocabulary** is specified, then both must be specified.

[PUBS-ID-00018][Error] If element **SubjectCode** has attribute **subjectCodeVocabulary** specified and does not have a value of "other", attribute **otherVocabulary** must not be specified.

[PUBS-ID-00019][Error] If element **SubjectCode** has attribute **subjectCodeVocabulary** with a value of "other," attribute **otherVocabulary** must be specified.

[PUBS-ID-00020][Error] If element **IntelDiscipline** has a value of "other," then attribute **otherDiscipline** must be specified.

[PUBS-ID-00021][Error] If element **IntelDiscipline** does not have a value of "other," then attribute **otherDiscipline** must not be specified.

[PUBS-ID-00022][Error] If element **PositiveIntel** has attribute **assertion** with a value of "other," then attribute **otherAssertion** must be specified.

[PUBS-ID-00023][Error] If element **PositiveIntel** has attribute **assertion** whose value is not "other," then attribute **otherAssertion** must not be specified.

[PUBS-ID-00024][Error] If element **Language has** attribute **encoding** with a value of "other," then attribute **otherVocabulary** must be specified.

[PUBS-ID-00025][Error] If element **Language has** attribute **encoding** whose value is not "other," then attribute **otherVocabulary** must not be specified.

[PUBS-ID-00026][Error] Element **IdentifierList**/**DocumentID** must exist and must have a non-null value.

[PUBS-ID-00027][Error] For element ImageArea attribute **coordinates** must be specified and the value must be appropriate to the value of attribute **shape** as defined by the PUBS.XML Data Element Dictionary.

[PUBS-ID-00028][Error] For elements **AnimationExhibit**, **AudioExhibit**, **Interactive3DExhibit**, **OtherExhibit**, **StillImageExhibit** and **VideoExhibit** attribute **xlink:href** must be specified.

[PUBS-ID-00029][Error] For elements **Facility**, **Person** and **TaskID** if attribute **xlink:href** exists, then the attribute must have a non-null value.

[PUBS-ID-00030][Error] For element **Glossary**, if child element **GlossaryDivision** is used, there must be at least two instances of element **GlossaryDivision**.

[PUBS-ID-00031][Error] For element **Bibliography**, if child element **BibliographyDivision** is used, there must be at least two instances of element **BibliographyDivision**.

[PUBS-ID-00032][Error] For element **Index**, if child element **IndexDivision** is used, there must be at least two instances of element **IndexDivision**.

[PUBS-ID-00065][Error] For elements:

- **Addressee**
- **AuthorInfo**
- **CoauthorInfo**
- **ContributingAuthorInfo**
- **POCinfo**
- **RequesterInfo**

At least one of the following elements must have non whitespace content

- **Surname**
- **UserID**
- **JobTitle**
- **Affiliation**
- **OfficeName**
- **PostalAddress**
- **PhoneNumber**
- **FaxNumber**
- **EmailAddress**
- **WebPageAddress**
- **FormattedSignatureBlock**

4.6.1 Elements ImageMap and StillImageExhibit

[PUBS-ID-00033][Error] Attribute **id** for element **ImageMap** must be specified and must have a unique value within the XML instance.

[PUBS-ID-00034][Error] For a given attribute **imageMapReference** value for element **StillImageExhibit**, there must be an attribute id for element **ImageMap** with an identical value within the XML instance.

4.6.2 Elements Link and ImageArea

[PUBS-ID-00035][Error] All links of attribute xlink:type simple must be potentially traversable. Attribute **xlink:href** must be specified.

[PUBS-ID-00036][Error] Internal links must be traversable. If the attribute **xlink:href** value for elements **Link** and **ImageArea** is a URL fragment identifier with a "#" prefix, that fragment identifier must match the attribute id value of an element within the XML instance. Note that a fragment identifier is characterized by a "#" prefix. The "#" prefix is not part of the fragment identifier. So when the **xlink:href** attribute value starts with a "#", the remainder of the value is the fragment identifier. This rule does not apply to attribute values which contain a URL and a fragment identifier separated by "#", in which case it can not be determined with certainty that the target is within the XML instance.

4.7 Attributes network and otherNetwork

[PUBS-ID-00037][Error] If attribute **network** is specified for an element and has a value of "other," then attribute **otherNetwork** must be specified for the same element and must have a value.

[PUBS-ID-00038][Error] If attribute **network** is specified for an element and does not have a value of "other," then attribute **otherNetwork** must not be specified for the same element.

[PUBS-ID-00039][Error] If attribute **network** is not specified for an element which allows the attribute, then attribute otherNetwork must not be specified for the same element.

4.8 Dates and Times

Except for attribute **date** for which the datatype is "xsd:date," the datatype of each PUBS.XML date/time-related element and attribute is specified as one of four custom simple types defined to allow the full range of allowable patterns specified in the DED for that element or attribute. These four custom simple types are in fact unions of appropriate W3C primitive datatypes, three of which also include in the union one of two additional custom simple types defined to allow seconds to be optional in time specifications. Schema validation will automatically ensure conformity to the datatypes. Validations and time comparisons will use the Zulu (Z) time zone when a time zone indicator is absent.

PUBS.XML constraint rules specify the following restrictions to the permissible values of dates and times:

[PUBS-ID-00040][Error] The permissible values for the year range are 1901 through the current year for elements **DateApproved**, **DateInfoCutoff**, **DatePosted**, **DatePublished**, **DateReviewed** and **DateRevised**.

[PUBS-ID-00041][Error] The permissible values for the year range are 1901 through 9999 for element **DateValidTil**.

[PUBS-ID-00042][Error] The permissible values for the year range are 0001 through 9999 for elements **DateEnd**, **DateStart**, **DateTimeReferenced** and **DateInformation**; and attributes **date**, **dateTime**, **normalizedDate** and **normalizedDateTime**.

[PUBS-ID-00043][Error] The permissible values for the decimal seconds are .0 through .999 for elements **DateApproved**, **DateEnd**, **DateInfoCutoff**, **DateInformation**, **DatePosted**, **DatePublished**, **DateReviewed**, **DateRevised**, **DateStart**, **DateTimeReferenced**, and **DateValidTil**; and attributes **dateTime**, **normalizedDateTime**, **normalizedTime**, and **dateTimeRange**.

[PUBS-ID-00044][Error] For attribute **dateTimeRange**, for each pair of date/time values, the second value must be later than the first value.

Table 8 summarizes the datatypes and allowable layout of representations for each of the PUBS.XML date/time-related elements and attributes.

Table 8. Date/Time-related DataTypes and Layout Representations

Element or Attribute	Datatype	Layout of Representation

Element or Attribute	Datatype	Layout of Representation
DateApproved **DateEnd** **DateInfoCutoff** **DatePosted** **DatePublished** **DateReviewed** **DateRevised** **DateStart** **DateValidTil** **@normalizedDateTime** **DateTimeReferenced** **DateInformation**	ISO8601DateTimeType	YYYY(Z | ±hh:mm)? | YYYY-MM(Z | ±hh:mm)? | YYYY-MM-DD(Z | ±hh:mm)? | YYYY-MM-DDThh:mm(Z | ±hh:mm) | YYYY-MM-DDThh:mm:ss(.s)?(Z | ±hh:mm)?
@date	xsd:date	YYYY-MM-DD(Z | ±hh:mm)?
@dateTime	dateTimesType	(YYYY(Z | ±hh:mm)? | YYYY-MM(Z | ±hh:mm)? | YYYY-MM-DD(Z | ±hh:mm)? | YYYY-MM-DDThh:mm(Z | ±hh:mm) | YYYY-MM-DDThh:mm:ss(.s)?(Z | ±hh:mm)?)+
@normalizedDate	ISO8601DateType	YYYY(Z | ±hh:mm)? | YYYY-MM(Z | ±hh:mm)? | YYYY-MM-DD(Z | ±hh:mm)?
@normalizedTime	ISO8601TimeType	hh:mm(Z | ±hh:mm) | hh:mm:ss(.s)?(Z | ±hh:mm)?
@dateTimeRange	dateTimePairsType	((YYYY(Z | ±hh:mm)? | YYYY-MM(Z | ±hh:mm)? | YYYY-MM-DD(Z | ±hh:mm)? | YYYY-MM-DDThh:mm(Z | ±hh:mm) | YYYY-MM-DDThh:mm:ss(.s)?(Z | ±hh:mm)?), (YYYY(Z | ±hh:mm)? | YYYY-MM(Z | ±hh:mm)? | YYYY-MM-DD(Z | ±hh:mm)? | YYYY-MM-DDThh:mm(Z | ±hh:mm) | YYYY-MM-DDThh:mm:ss(.s)?(Z | ±hh:mm)?))+

4.9 Cross Element and Attribute Constraint Rules

A PUBS.XML document must conform to the following constraint rules.

[PUBS-ID-00045][Warning] **DateApproved** must not be later than **DatePublished**, and **DatePosted**.

[PUBS-ID-00046][Error] **DateInfoCutoff** must not be later than **DatePublished**, and **DatePosted**.

[PUBS-ID-00047][Warning] **DateValidTil** must not be earlier than **DateApproved**, **DateInfoCutoff**, **DatePublished**, and **DatePosted**.

[PUBS-ID-00048][Warning] **DateRevised** must not be earlier than **DateApproved**, **DateInfoCutoff**, **DatePublished**, and **DatePosted**.

[PUBS-ID-00049][Error] **DateRevised** must not be later than **DateValidTil**

[PUBS-ID-00050][Warning] **DateReviewed** must not be later than **DateApproved**, **DatePublished**, and **DatePosted**.

[PUBS-ID-00051][Error] For any element **TimePeriod**, if elements **DateStart** and **DateEnd** are both specified as child elements, **DateStart** must not be later than **DateEnd**.

[PUBS-ID-00052][Error] For any element **TimePeriod**, an attribute **periodName** or a child element **DateStart** or **DateEnd** must be specified.

[PUBS-ID-00053][Error] For any element **PostalAddress**, at least one of its child elements **AddressLine**, **City**, **State**, **Province**, **PostalCode**, or **CountryCode** must have element content.

4.10 Time Zone Indicators

Validations and time comparisons will use the Zulu (Z) time zone, when a time zone is absent. It is recommended that the optional time zone be specified either as Zulu (Z) or as ±hh:mm where applicable.

[PUBS-ID-00054][Warning] For elements **DateApproved**, **DateEnd**, **DateInfoCutoff**, **DatePosted**, **DatePublished**, **DateReviewed**, **DateRevised**, **DateStart**, **DateValidTil**, **DateTimeReferenced**, and **DateInformation**, if the time designator (T) is specified, it is recommended that time zone be specified.

[PUBS-ID-00055][Warning] For attributes **dateTime**, **normalizedDateTime**, and **normalizedTime**, it is recommended that time zone be specified.

4.11 Information Security Markings (ISM.XML)

Constraint rules specific to the application of information security markings are documented in the *XML Data Encoding Specification for Information Security Marking Metadata* and related documents.

Chapter 5 – Data Rendering Constraint Rules

The constraint rules in this chapter define constraints of the rendering of PUBS.XML documents. The intent of this section is to inform the development of systems rendering or displaying PUBS.XML data for use by individuals not familiar with the details of the PUBS.XML markup. While expressed in a similar manner to the data validation constraint rules above, there is no expectation that evaluation of these rules can be automated; rather these rules should inform the evaluation of a system's capabilities and functionality.

5.1 Rule Identifiers

Each constraint rule has an assigned rule ID, indicated in brackets preceding the constraint rule description. The rule ID's from 00001 to 10000 are unclassified and10001 to 20000 are "for official use only" (FOUO). IDs from 20001 to 30000 are reserved for Secret rules and 30001 and above for more classified rules. PUBS.XML data rendering constrain rule IDs are prefixed with "PUBS-RENDER-"

As the constraint rules are managed over time, IDs from deleted rules will not be reused.

5.2 Errors and Warnings

The severity of a constraint rule violation is categorized as either an "Error" or a "Warning" and is indicated in brackets preceding each constraint rule description. An "Error" is more severe and is indicative of a clear violation of a constraint rule, which would be likely to have a significant impact on the quality of a system. A "Warning" is less severe although noteworthy, and may not necessarily have any impact on the quality of a system.

Each system responsible for rendering documents must be evaluated based on its use. Those evaluating the system must make a mission-appropriate decision about the system's suitability for use.

[PUBS-RENDER-00001][Error] When element **RevisionRecall** is present the text of the attribute **revisionType** shall be rendered in uppercase as the first part of the document title immediately following the classification portion mark and will be followed with a ":".
Human Readable: Systems used for rendering data containing the RevisionRecall element will produce rendered documents that comply with the August 5, 2005 Negroponte Revision Recall Memo, and the specific style described.

Appendix A IC CIO Approval Memo

An Office of the Intelligence Community Chief Information Officer (OCIO) Approval Memo should accompany this enterprise technical data specification bearing the signature of the Intelligence Community Chief Information Officer (IC CIO) or an IC CIO-designated official(s). If an OCIO Approval Memo is not accompanying this specification's version release package, then refer back to the authoritative web location(s) for this specification to see if a more complete package or a specification update is available.

Specification artifacts display a date representing the last time a version's artifacts as a whole were modified. This date most often represents the conclusion of the IC Element collaboration and coordination process. Once the IC Element coordination process is complete, the specification goes through an internal OCIO staffing and coordination process leading to signature of the OCIO Approval Memo. The signature date of the OCIO Approval Memo will be later than the last modified date shown on the specification artifacts by an indeterminable time period.

Upon signature of the OCIO Approval Memo, IC Elements may begin to use this specification version in order to address mission and business objectives. However, it is critical for IC Elements, prior to disseminating information encoded with this new specification version, to ensure that key enterprise services and consumers are prepared to accept this information. IC Elements should work with enterprise service providers and consumers to orchestrate an orderly implementation transition to this specification version in concert with mandatory and retirement usage decisions captured in the IC Enterprise Standards Baseline as defined in Intelligence Community Standard (ICS) 500-20.

Appendix B Acronyms

CAPCO – Controlled Access Program Coordination Office

DCMI – Dublin Core Metadata Initiative

DC MES – Dublin Core Metadata Element Set

DES – Data Encoding Specification

DOI – Digital Object Identifier

GNS – Geographic Names Server

HTML – HyperText Markup Language

IC CIO – Intelligence Community Chief Information Officer

ICD – Intelligence Community Directive

ICS – Intelligence Community Standard

ISBN – International Standard Book Number

ISO – International Organization for Standardization

ISOO – Information Security Oversight Office

KA – Knowledge Assertion

KOS – Knowledge Organization System

MIME – Internet Media Types

NARA – National Archives and Records Administration

NGA – National Geospatial Intelligence Agency

ODNI – Office of the Director of National Intelligence

TGN – Thesaurus of Geographic Names

URI – Uniform Resource Identifier

URL – Uniform Resource Locator

W3CDTF – World Wide Web Consortium Date Time Format

XML – Extensible Markup Language

Appendix C Glossary

No pertinent glossary items requiring further definition.

Appendix D Bibliography

(CAPCO Implementation Guide)
Intelligence Community Classification and Control Markings Implementation Manual. Unclassified FOUO version. Vol. 2, Ed. 2 (Version 2.2). 25 September 2009. Director of National Intelligence (DNI), Special Security Center (SSC), Controlled Access Program Coordination Office (CAPCO). https://www.intelink.gov/sites/ssc/capco/Shared%20Documents/CAPCO%20Implementation%20Manual_FOUO_v2.2_25Sep2009.pdf

(CAPCO Register)
Authorized Classification and Control Markings Register. Unclassified FOUO version. Vol. 2, Ed. 2 (Version 2.2). 25 September 2009. Director of National Intelligence (DNI), Special Security Center (SSC), Controlled Access Program Coordination Office (CAPCO). https://www.intelink.gov/sites/ssc/capco/Shared%20Documents/CAPCO%20Register_FOUO_2.2_25Sep2009.pdf

(DC MES)
Dublin Core Metadata Element Set. Version 1.1. 02 June 2003. Dublin Core Metadata Initiative. http://dublincore.org/documents/dces/.

(E.O. 12958, as amended)
Executive Order 12958 – Classified National Security Information, as Amended. Federal Register, Vol. 68, No. 60. 25 March 2003. The White House. http://www.archives.gov/isoo/policy-documents/eo-12958-amendment.html.

(E.O. 12829, as amended)
Executive Order 12958 – National Industrial Security Program, as Amended. Federal Register, Vol. 58, No. 240. 16 December 1993. The White House. http://www.archives.gov/isoo/policy-documents/eo-12829.html

(ICS 2007-500-3)
Intelligence Community Standard for Information Resource Metadata. Intelligence Community Standard Number 2007-500-3. 11 December 2007. Office of the Director of National Intelligence. http://www.dni.gov/ICIS/dsca/irm/ICS2007-500-3ResourceMetadata.pdf.

(ICS 2007-500-4)
Intelligence Community Standard for Publication Metadata. Intelligence Community Standard Number 2007-500-4. 11 December 2007. Office of the Director of National Intelligence. http://www.dni.gov/ICIS/dsca/pubs/ICS2007-500-4PublicationMetadata.pdf.

(ICS 2008-500-2)
Intelligence Community Standard for Source Reference Citation Metadata. Intelligence Community Standard Number 2008-500-2. 26 June 2008. Office of the Director of National Intelligence. http://www.dni.gov/ICIS/dsca/sc/ICS2008-500-2SourceReferenceCitationMetadata.pdf.

(ICS 2008-500-7)
Intelligence Community Standard for Information Security Marking Metadata. Intelligence Community Standard Number 2008-500-7. 19 August 2008. Office of the Director of National Intelligence. http://www.dni.gov/ICIS/dsca/ism/ICS2008-500-7SecurityMarkingMetadata.pdf.

(ICD 206)
Sourcing Requirements for Disseminated Intelligence Products. Intelligence Community Directive Number 206. 17 October 2007. Office of the Director of National Intelligence. http://www.dni.gov/electronic_reading_room/ICD_206.pdf.

(ICD 500)
Director of National Intelligence Chief Information Officer. Intelligence Community Directive Number 500. 7 August 2008. Office of the Director of National Intelligence. http://www.dni.gov/electronic_reading_room/ICD_500.pdf.

(ISO 639-2)
Codes for the representation of names of languages – Part 2: Alpha-3 code. ISO 639-2:1998. International Organization for Standardization (ISO). http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=4767.

(ISO 3166-1)
Codes for the representation of names of countries and their subdivisions – Part 1: Country codes. ISO 3166-1:2006. International Organization for Standardization (ISO). http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39719.

(ISO 8601)
Data elements and interchange formats – Information interchange – Representation of dates and times. ISO 8601:2004. International Organization for Standardization (ISO). http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40874.

(ISO 15836)
Information and documentation – The Dublin Core metadata element set. ISO 15836:2009. International Organization for Standardization (ISO). http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=52142.

(ISOO Directive 1)
Classified National Security Information (Directive No. 1); Final Rule. 32 CFR Parts 2001 and 2004. Federal Register, Vol. 68, No. 183. 22 September 2003. Information Security Oversight Office (ISOO), National Archives and Records Administration (NARA). http://www.archives.gov/isoo/policy-documents/eo-12958-implementing-directive.pdf.

(RFC 3066)
Tags for the Identification of Languages. January 2001. H. Alvestrand. Cisco Systems. http://www.rfc-editor.org/rfc/rfc3066.txt.

Appendix E Points of Contact

This technical specification is managed by the Office of the Intelligence Community Chief Information Officer (IC CIO). As of this writing, the IC CIO/IC Enterprise Architecture (ICEA) Directorate facilitates the IC data collaboration and coordination forums responsible for the selection or development of common IC technical data specifications. Direct all inquiries about this IC technical specification to IC CIO/ICEA, the IC's data collaboration and coordination forum, or IC element representatives involved in those forums.

Appendix F Change History

The following table summarizes the version identifier history for this DES.

Table 9. DES Version Identifier History

Version	Identifier	Date	Purpose
1.0	None	August 2008	Initial Release
2	ICTechSpec 500.D.3-V2	24 December 2009	Routine revision to technical specification. For details of changes, see 1.3.

Appendix G Configuration Management

The selection or development of technical data specifications of common interest to the IC are collaborated and coordinated currently within governance forums managed by the IC CIO. Change requests for this technical data specification should be directed to the office identified in Appendix E – Points of Contact.

Appendix H Reading the Schematics

The physical XML structures documented in this guide are illustrated with schematics created with the commercially available product named *XML Spy®,* produced by Altova. The symbology used by *XML Spy®* is described in this appendix.

H.1 Element Models

In *XML Spy*, XML elements are represented by rectangles like that shown for an element named **PublicationMetadata** in Figure 1. XML elements may have one of three types of content model:

- Element content: A model in which the content consists entirely of child elements; in other words, there is no *direct* data content (although the child elements may have data content).
- Mixed content: A model in which the content consists of text, possibly intermixed with child elements; child elements in mixed content are said to float in the text in that their use is not constrained to a hierarchy.
- Empty: A model in which the element has no content; the element's function is performed by its attribute(s), if any, or the element is simply a placeholder for data that will be generated when the element is rendered for presentation.

Figure 1. Symbol for an XML Element having Element Content

An element whose model is element content is illustrated by the preceding figure. The "+" symbol in the small box at the right edge denotes that **PublicationMetadata** has content.

In the *XML Spy* graphical user interface (GUI), a schematic may be expanded and contracted to expose more or less of the model. In the GUI, clicking with the mouse on the "+" symbol causes the content of the element (in this case **PublicationMetadata**) to be revealed and the "+" is replaced by a "-" symbol. The purpose for pointing this out is that, in the remainder of this appendix, some of the illustrations show the "-" rather than the "+".

An element whose model is *mixed content* is represented by the symbol shown in Figure 2. As above, the "+" signifies that the element has content, and the lines in the upper left corner denote that text is allowed. Mixed content is the normal model for

elements like paragraphs, list items, titles, *et al.*, in which the text "contains" semantic objects like footnote references, superscripts and subscripts, italicized and bolded passages, quoted strings, words or phrases that have been tagged for indexing or searching, *etc.* These types of semantic objects that appear in the running text are represented in XML as elements.

Figure 2. Symbol of an Element with Mixed Content

An *empty* element is illustrated by Figure 3. Note that the element symbol lacks a "+" symbol, meaning that content is not allowed. Empty elements do not have textual content. They are used for one or more of three possible purposes:

- They are pointers to text or non-text objects that are located in external files. In this capacity, they can be used as placeholders for illustrations. They mark the position where a graphic is to be placed and, using attributes, identify the file that contains the graphic image.
- They are placeholders for application-generated content that will be created as part of the rendering process. Examples include tables of contents, indexes, glossaries, and bibliographies. Empty elements are used to specify the positioning of such constructs.
- They mark locations that are acted upon by formatting engines. A common example is to use an empty element called **br** or **break** to tell a formatter to insert a newline sequence at the point where the empty element is encountered.

Figure 3. Symbol of an Empty Element

H.2 Groupings of Elements

XML Spy represents an XML schema element group that has been used in a content model with a rounded box. In Figure 4, **ComplexContentGroup** is the name of an element group.

Figure 4. Element Group Symbol

Note the use of the "+" symbol. In the GUI, clicking on this symbol reveals the composition of the element group. This is illustrated later on in this appendix.

H.3 Occurrence indicators

Some elements must be used, they are designated as required. Other elements are designated as optional. Both required and optional elements may be repeatable. These in combination make four classes of occurrence indicators: a required element that can be used once and only once, a required element that may repeat, an optional and repeatable element, and an optional element that may not repeat.

Figure 5 illustrates several elements the use of which is optional. The dotted border of the element symbols for **UUID**, **DocumentID**, and **InternalID** denote optionality. This means that the producer may use these elements to tag document identifiers, but their use is not required. The XML document will be deemed *syntactically* correct with or without them.

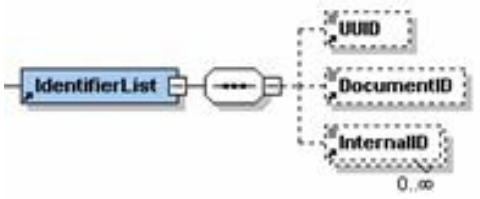

Figure 5. Optional elements, one repeatable, graphical representation

It is important to emphasize, in conjunction with Figure 5, that syntactic correctness and semantic completeness are distinct requirements. While an automated XML processing software application will accept element **IdentifierList** with or without the child elements, a production organization will need to apply local business rules to require at least one of the child elements to be present. This is done in order to allow the model to be used throughout the life cycle of the data. For example, the model can be used both at the time of authoring when the values for the child elements may not be known and later on when the values are required during the exchange with a partner.

The next two figures illustrate the symbols for repeatable elements. In Figure 6, element **InternalID** may be used zero or more times, as indicated by the "0..∞" string beneath the symbol. Note also that a repeatable element is denoted by a stack of element symbols.

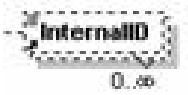

Figure 6. Optional and repeatable element, graphical representation

By contrast, in Figure 7, the element named **Section** must be used at least once, and may be used multiple times in succession. The fact that the section element symbol has solid borders means that its use is required. The "1..∞" beneath the symbol means it can be used one or more times.

Figure 7. Required and repeatable element, graphical representation

An element symbol that has solid borders and neither "0..∞" nor "1..∞" beneath must be used one and only one time. In Figure 5, the element named **IdentifierList** must be used once and only once.

H.4 Order of Elements

In many cases, XML schemas require that elements be used in a prescribed sequence. As shown in Figure 8 this constraint is denoted by an ellipsis in a rounded box in the schematics:

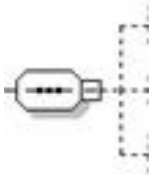

Figure 8. Sequence Symbol

The connecting lines leading from the right edge lead to three elements which, if used, must be used in top-to-bottom order. Figure 5 shows this symbol in a larger context. It says that the subelements of **IdentifierList** are all optional but, if used, must be in the order shown.

Figure 9 shows the symbol that signifies that a choice must be made between the child components. The symbol is a three-pole switch inside of a rounded box. The switch signifies that an author can choose either of the branches leading from the right edge.

Because the borders of the box are dashed, the choice itself is optional. The "0..∞" symbol beneath the box means that the choice can be made multiple times.

Figure 9. Choice Symbol

Figure 10 shows the use of this symbol in context:

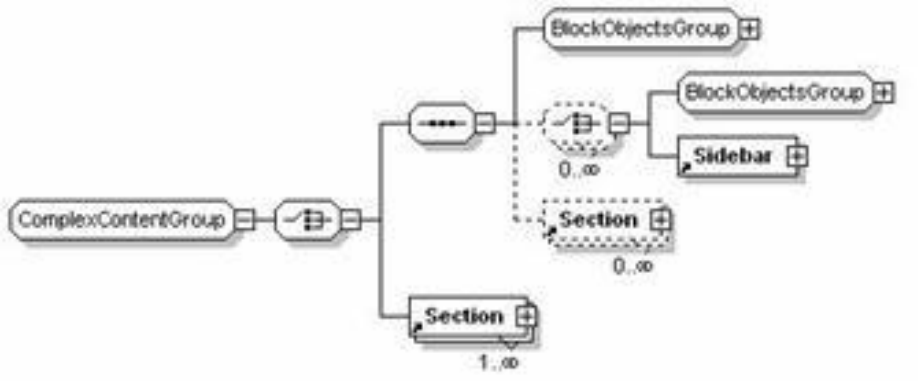

Figure 10. Sequence and Choice Combined

This schematic says that the grouping named **ComplexContentGroup** consists of a required choice between the top branch on the one hand and the bottom branch on the other. The top branch is a sequence of one element from the **BlockObjectsGroup**, optionally followed by a *repeatable* choice between **BlockObjectsGroup** elements and a **Sidebar** element, optionally followed by zero or more **Section** elements. The bottom branch leads to one or more **Section** elements.

H.5 Attributes

Elements may have properties. For example, the element **List** has the property called **listStyle**. In XML, such properties of an element are called *attributes* of the element. In the schematic diagrams used in this document, the attributes are displayed separately from the elements.

Appendix I Controlled Vocabulary Enumerations

The Controlled Vocabulary Enumerations (CVE) used in this DES are as follows:

Table 10. CVE Definitions

CVE File name	Definition	Attribute and Rules Cross reference
CVEnumPubsMimeType	All currently valid MIME Types.	@MIMEtype PUBS-ID-00074 PUBS-ID-00075
CVEnumUNCE20UnitsOfMeasure	All currently valid UNCE-20 units of measure.	@unitOfMeasure PUBS-ID-00062
CVEnumPubsSubjectCodeSystems	All currently valid Subject Code Systems.	@subjectCodeVocabulary SchemaEnumeration
CVEnumPubsIntelSubDisciplineTechniques	All currently valid Intel sub Discipline technique codes.	IntelSubdisciplineTechnique SchemaEnumeration
CVEnumPubsIntelSubDisciplines	All currently valid Intel sub Discipline codes.	IntelSubdiscipline SchemaEnumeration
CVEnumPubsIntelDisciplines	All currently valid Intel Discipline codes.	IntelDiscipline SchemaEnumeration
CVEnumPubsCoverageISO3166Trigraph	WWW and All currently valid ISO-3166 trigraphs	CountryCode PUBS-ID-00068
CVEnumISO639-2Trigraph	All currently valid ISO-639-2 Trigraphs.	Language PUBS-ID-00057 PUBS-ID-00060
CVEnumPubsCoverageISO3166Numeric	All currently valid ISO-3166 numeric values	CountryCode PUBS-ID-00069
CVEnumPubsCoverageISO3166Digraph	All currently valid ISO-3166 diagraphs.	CountryCode PUBS-ID-00067 PUBS-ID-00059 PUBS-ID-00060 PUBS-ID-00061
CVEnumPubsCoverageFIPSDigraph	All currently valid FIPS diagraphs.	CountryCode PUBS-ID-00066 PUBS-ID-00011
CVEnumPubsCountryISO3166Trigraph	All ISO-3166 trigraphs that have ever been valid	@countryCode PUBS-ID-00072
CVEnumPubsCountryISO3166Numeric	All ISO-3166 numeric values that have ever been valid	@countryCode PUBS-ID-00073
CVEnumPubsCountryFIPSDigraph	All FIPS 10-4 digraphs that have ever been valid	@countryCode PUBS-ID-00071
CVEnumPubsAgencyAcronym	All currently valid Agency Acronyms for use with publisher.	AgencyAcronym PUBS-ID-00076
CVEnumISO4217Trigraph	All currently valid ISO-4217 Trigraphs.	@unitOfMeasure PUBS-ID-00063
CVEnumISO4217Numeric	All currently valid ISO-4217 Numeric codes.	@unitOfMeasure PUBS-ID-00064

CVE File name	Definition	Attribute and Rules Cross reference
CVEnumISO639Digraph	All currently valid ISO-639-1 Digraphs.	Language PUBS-ID-00056 PUBS-ID-00059 PUBS-ID-00061
CVEnumISO639-3Trigraph	All currently valid ISO-639-3 Trigraphs.	Language PUBS-ID-00058

Appendix J IRM-HTML to PUBS.XML Mapping Information

The "*Intelligence Community Information Resource Metadata, HTML Encoding*" is used by many producing organizations to encapsulate metadata in HTML documents posted to Intelink. In the HTML Implementation, the metadata is stored in **META** elements within the HTML **HEAD** element. This is illustrated by the following example:

```
<HTML>
<HEAD>
<TITLE>(U) Example for Intelink Posting (UNCLASSIFIED)</TITLE>
<META NAME="IL.secur.classif" CONTENT="UNCLASSIFIED">
<META NAME="IL.title" CONTENT="(U) Example for Intelink Posting">
<META NAME="IL.docid" CONTENT="D192-012">
<MEAT NAME="IL.country" CONTENT="AFG">
<META NAME="IL.subcode.ifc" CONTENT="1011">
...
<META NAME="IL.summary" CONTENT="(U) Abstract of this
document...">
</HEAD>
<BODY>
...
</BODY>
</HTML>
```

The HTML Implementation mandates security, product, date, terms-of-use, and electronic (*i.e.*, media format) metadata. The same metadata required by the HTML Implementation is also required by PUBS-XML.

In PUBS-XML, the security metadata is handled by the IC ISM attributes of the element **Security** contained in **PublicationMetadata**. For the other metadata categories individual XML elements are provided. For example, "IL.title" corresponds to element **Title** when used in the context that its parent is element **DescriptiveMetadata** and its ancestor is **PublicationMetadata**. Those mapping starting with element **PublicationMetadata** are, in turn, the first child of one of the seven a document's top-level element (such as **AnalyticalPacket, Article, Basic,** *et al.*). When metadata from the HTML Implementation maps to an attribute in PUBS-XML, the attribute name, preceeded by an "@", is appended to the path. The following path addresses attribute **SCIcontrols** of element **Security**:
PublicationMetadata/DescriptiveMetadata/Security/@SCIcontrols

This appendix contains a mapping of the metadata elements from the HTML Implementation to the corresponding elements and attributes in the PUBS-XML

implementation. This is a list of all the actual XML "tags" (XML Element and Attribute *names*) that can appear in an XML document file that maps the Information Resource Metadata (HTML Encoding) elements to its corresponding Intelligence Community (PUBS-XML) Mapped elements and/or attribute. The PUBS-XML Schema essentially defines a tree structure, containing distinct nodes, each of which can have zero, one, or many associated atomic data values (within an actual schema instance). To explain these mandatory-versus-optional, and singular-versus-repeating constraints, one of the following supplementary annotations has been appended to each tag name.

Table 11. Occurrence codes used in mapping table

Code	Comments
[MS]	Mandatory, Singular (attached to non-repeating, non-optional Elements; also attached to Attributes marked use="required");
[MR]	Mandatory, Repeating (attached to Elements marked maxOccurs="unbounded"; also attached to "list" Attributes, like ownerProducer, which is of type NMTOKENS)
[OS]	Optional, Singular (attached to Elements marked minOccurs="0", but not maxOccurs="unbounded"; also attached to optional Attributes, such as those marked use="optional", but not base="xsd:NMTOKENS")
[OR]	Optional, Repeating (attached to Elements marked minOccurs="0" and maxOccurs="unbounded"; also attached to optional "list" Attributes, such as those marked use="optional" and base="xsd:NMTOKENS"; also attached----rather than using perhaps [XOR]----to any Elements that are listed in an XSD clause of the form: <xs:choice minOccurs="0" maxOccurs="unbounded">, since the listed Elements are neither required nor mutually exclusive in this case)
[MSX]	Mandatory, Singular, but mutually eXclusive with its peers (attached to any Elements that are listed in an XSD clause of the form: <xs:choice>, because these choice structures are subject to the default restrictions minOccurs="1" and maxOccurs="1")
Note that XML Schema notation enables repetition to be accommodated in a wide variety of ways. Our interest here is in the existence of repetition of any kind, due to its profound affect on XML valuation mappings for catalog harvest/ingest operations.	

Table 12. IRM.HTML to PUBS.XML XPath Mappings

IRM-HTML	PUBS-XML XPath Mapping
Security Metadata	
IL.secur.classif	PublicationMetadata/DescriptiveMetadata/Security/@classification [MS]
IL.secur.ownerproducer	PublicationMetadata/DescriptiveMetadata/Security/@ownerProducer [MR]
IL.secur.ctrl	PublicationMetadata/DescriptiveMetadata/Security/@SCIcontrols [OR]
IL.secur.saridentifier	PublicationMetadata/DescriptiveMetadata/Security/@SARIdentifier [OR]
IL.secur.fgi	PublicationMetadata/DescriptiveMetadata/Security/@FGIsourceOpen [OR] PublicationMetadata/DescriptiveMetadata/Security/@FGIsourceProtected [OR]
IL.secur.dissem	PublicationMetadata/DescriptiveMetadata/Security/@disseminationControls [OR]
IL.secur.relto	PublicationMetadata/DescriptiveMetadata/Security/@releasableTo [OR]
IL.secur.nonic	PublicationMetadata/DescriptiveMetadata/Security/@nonICmarkings [OR]
IL.secur.declassmanualreview	No mapping

IRM-HTML	PUBS-XML XPath Mapping
IL.secur.declasson	PublicationMetadata/DescriptiveMetadata/Security/@declassDate [OS]
Product Metadata	
IL.title	PublicationMetadata/DescriptiveMetadata/Title [MS]
IL.docid	PublicationMetadata/AdministrativeMetadata/IdentifierList/DocumentID [OS]
IL.summary	PublicationMetadata/DescriptiveMetadata/Description [MS]
IL.keyword	PublicationMetadata/DescriptiveMetadata/Subject/Keyword [MR]
IL.country	Publisher/AuthorInfo/PostalAddress/CountryCode [OS]
IL.subcode.xxx	PublicationMetadata/DescriptiveMetadata/Subject/SubjectCode [OR]
IL.agency	PublicationMetadata/AdministrativeMetadata/Publisher/AgencyAcronym [MS]
IL.poc	PublicationMetadata/AdministrativeMetadata/Publisher/POCInfo [OR]
IL.itype	PublicationMetadata/DescriptiveMetadata/IntelType [OR]
IL.productline	PublicationMetadata/DescriptiveMetadata/ProductLine [OS]
Date Metadata	
IL.pubdate	PublicationMetadata/AdministrativeMetadata/DateList/DatePublished [MS]
IL.postdate	PublicationMetadata/AdministrativeMetadata/DateList/DatePosted [MSX]
IL.cutdate	PublicationMetadata/AdministrativeMetadata/DateList/DateInfoCutoff [OS]
IL.validtil	PublicationMetadata/AdministrativeMetadata/DateList/DateValidTil [OS]
Terms of Use Metadata	
IL.privacyact	PublicationMetadata/AdministrativeMetadata/Rights/PrivacyActIndicator [MS]
IL.vitalrec	PublicationMetadata/AdministrativeMetadata/RecordsManagementInfo/VitalRecordIndicator [MS]
IL.copyright	PublicationMetadata/AdministrativeMetadata/Rights/CopyRightIndicator [MS]
Electronic Metadata	
IL.applicationtitle	PublicationMetadata/AdministrativeMetadata/RecordsManagementInfo/ApplicationSoftware [MS]
IL.format	PublicationMetadata/DescriptiveMetadata/Description/MediaResource [MS]
IL.url	No Mapping
IL.itype.subdiscipline	PublicationMetadata/DescriptiveMetadata/IntelType/IntelSubdiscipline [OS]
IL.requirement	No Mapping
IL.sensor	No Mapping
IL.reportphase	No Mapping
IL.datetime.range	PublicationMetadata/DescriptiveMetadata/Coverage/Temporal/TimePeriod [MR]
IL.datetime.point	No Mapping
IL.analysistool	No Mapping
IL.designatedactivity	No Mapping